TECHNICAL NOTES

DATA SOURCES

The California Department of Health Services (CDHS), Center for Health Statistics, Office of Health Information and Research, was the source for the data that appear in this report. Numerator data were extracted from the Death Statistical Master Files for the years 2000-2002, from the Multiple Cause of Death Files for 2000-2001, from the Birth Statistical Master Files for 2000-2003, from the Birth Cohort Perinatal Files for 2000-2001, from the Office of Statewide Health Planning and Development (OSHPD) Patient Discharge Data files for 2000-2002, and from the California Health Interview Survey (AskCHIS) query system for 2001.[1-6] Denominator data by age, sex, and race-ethnicity used in the calculation of population-based mortality and morbidity rates were supplied by the California Department of Finance (CDOF), Demographic Research Unit (DRU).[7] Denominator data used in the calculation of natality rates were extracted from the Birth Statistical Master Files and from the Birth Cohort Perinatal Files.

DATA DEFINITIONS

Race and Ethnicity

Numerator data (i.e., the number of births or deaths used in the calculation of mortality or natality rates) were extracted from the birth and death files using the "Hispanic Origin" and "Multi-Race" variables, which provide the following mutually exclusive categories:

Hispanic/Latino (Hispanic Origin codes 2-8)

African American/Black, non-Hispanic (Multi-Race code 2; Hispanic Origin codes 1,9)

American Indian/Alaska Native, non-Hispanic (Multi-Race code 3; Hispanic Origin codes 1,9)

Asian, non-Hispanic (Multi-Race codes 4; Hispanic Origin codes 1,9)

Native Hawaiian/Other Pacific Islander (Multi-Race code 5; Hispanic Origin codes 1,9)

White, non-Hispanic (Multi-Race code 1; Hispanic Origin codes 1,9)

Multirace, non-Hispanic (Multi-Race code 7; Hispanic Origin codes 1,9)

Other/Unknown, non-Hispanic (Multi-Race codes 6,9; Hispanic Origin codes 1,9)

Numerator data (i.e., the number of hospitalizations used in the calculation of morbidity rates) were extracted using the "Ethnicity" and "Race" variables available in the OSHPD PDD, which provide these mutually exclusive categories:

Hispanic/Latino (Ethnicity code 1)

African American/Black, non-Hispanic (Race code 2; Ethnicity codes 2,3)

American Indian/Alaska Native, non-Hispanic (Race code 3; Ethnicity codes 2,3)

Asian/Pacific Islander, non-Hispanic (Race code 4; Ethnicity codes 2,3)

White, non-Hispanic (Race code 1; Ethnicity codes 2,3)

Other/Unknown, non-Hispanic (Race codes 5,6; Ethnicity codes 2,3)

Denominator data used in the calculation of population-based rates were extracted from the CDOF/DRU:

Total (All Racial and Ethnic Groups), by Sex, by Age Group
African American/Black, non-Hispanic, by Sex, By Age Group
American Indian/Alaska Native, non-Hispanic, by Sex, by Age Group
Asian/Pacific Islander, non-Hispanic, by Sex, by Age Group
  Asian, non-Hispanic, by Sex, by Age Group
  Native Hawaiian/Other Pacific Islander, by Sex, by Age Group
Hispanic/Latino, by Sex, by Age Group
White, non-Hispanic, by Sex, by Age Group
Multirace, non-Hispanic, by Sex, by Age Group

Mortality Data

Beginning in 1999, deaths are coded using the International Classification of Diseases, Tenth Revision (ICD-10).[8] Readers and users of these data are cautioned that prior year's mortality tables based on ICD-9 are not comparable and should not be used for direct comparisons with ICD-10 data.

Following is a list of the ICD-10 codes that were used to operationally define each Healthy People 2010 objective that uses mortality data for monitoring:[9-10]

Obj. 3-1	All Cancer Deaths	C00-C97
Obj. 3-2	Lung Cancer Deaths	C33-C34
Obj. 3-3	Female Breast Cancer Deaths	C50
Obj. 3-4	Cervical Cancer Deaths	C53
Obj. 3-5	Colorectal Cancer Deaths	C18-C21
Obj. 3-6	Oropharyngeal Cancer Deaths	C00-C14
Obj. 3-7	Male Prostate Cancer Deaths	C61
Obj. 3-8	Malignant Melanoma Deaths	C43
Obj. 5-5	Diabetes-Related Deaths	E10-E14
Obj. 12-1	Coronary Heart Disease Deaths	I11, I20-I25
Obj. 12-7	Stroke Deaths	I60-I69
Obj. 13-14	HIV Deaths	B20-B24
Obj. 15-3	Firearm-Related Deaths	W32-W34, X72-X74, X93-X95, Y22-Y24, Y35.0
Obj. 15-8	Poisoning Deaths	X40-X49, X60-X69, X85-X90, Y10-Y19, Y35.2
Obj. 15-9	Suffocation Deaths	X70, X91, W75-W84, Y20
Obj. 15-13	Unintentional Injury Deaths	V01-X59, Y85-Y86
Obj. 15-15	Motor Vehicle Crash Deaths	V30-V39 (.4-.9), V40-V49 (.4-.9), V50-V59 (.4-.9), V60-V69 (.4-.9), V70-V79 (.4-.9), V81.1, V82.1, V83-V86 (.0-.3), V20-V28 (.3-.9), V19 (.4-.6), V29 (.4-.9), V12-V14 (.3-.9),

Obj. 15-15, cont'd		V02-V04 (.1-.9), V09.2, V80 (.3-.5), V87 (.0-.8), V89.2
Obj. 15-16	Pedestrian Deaths	V02-V04 (.1, .9), V09.2
Obj. 15-25	Residential Fire Deaths	X00-X09
Obj. 15-27	Fall-Related Deaths	W00-W19
Obj. 15-29	Drowning Deaths	W65-W74, V90, V92
Obj. 15-32	Homicide Deaths	X85-Y09, Y87.1
Obj. 16-4	Maternal Deaths Due To Obstetric Causes	O00-O95, O98-O99
Obj. 18-1	Suicide Deaths	X60-X84, Y87.0
Obj. 24-1	Asthma Deaths	J45-J46
Obj. 24-10	Chronic Obstructive Pulmonary Disease Deaths	J40-J47
Obj. 26-2	Cirrhosis Deaths	K70, K73, K74
Obj. 26-3	Drug-Induced Deaths	F11.0-F11.5, F11.7-F11.9,F12.0-F12.5, F12.7-F12.9, F13.0-F13.5, F13.7-F13.9, F14.0-F14.5, F14.7-F14.9, F15.0-F15.5, F15.7-F15.9, F16.0-F16.5, F16.7-F16.9, F17.0-F17.5, F17.7-F17.9, F18.0-F18.5, F18.7-F18.9, F19.0-F19.5, F19.7-F19.9, X40-X44, X60-X64, X85, Y10-Y14

Infant Mortality

The infant death rate is the number of deaths among infants under one year of age per 1,000 live births, and is a universally accepted and easily understood indicator that is useful in pointing to problems with the health status of infants and mothers, and possible problems in the delivery of health care and related services to these groups in a community. The following HP2010 objectives use data from the linked birth and infant death records available in the California Birth Cohort Perinatal Outcome Files for 2000-2001:

Obj. 26-2	Infant Deaths	A00-Y89
Obj. 16-1f	Birth Defects Deaths	Q00-Q99
Obj. 16-1g	Congenital Heart and Vascular Defects Deaths	Q20-Q28
Obj. 16-1h	Sudden Infant Death Syndrome Deaths	R95

Child, Adolescent, and Young Adult Mortality

The following HP2010 objectives use data from the Death Statistical Master Files for 2000-2002:

Obj. 16-2a	Child Deaths (Ages 1-4)	A00-Y89
Obj. 16-2b	Child Deaths (Ages 5-9)	A00-Y89
Obj. 16-3a	Adolescent Deaths (Ages 10-14)	A00-Y89
Obj. 16-3b	Adolescent Deaths (Ages 15-19)	A00-Y89
Obj. 16-3c	Young Adult Deaths (Ages 20-24)	A00-Y89

Natality

The following HP2010 objectives use natality data obtained from the Birth Statistical Master Files for 2000- 2003:

Obj. 16-6a The prenatal care indicator, Month Prenatal Care Began, has been associated with access to care. Early prenatal care is defined as the percentage of mothers who began prenatal care in their first trimester (the first three months of pregnancy). However, as a health indicator, the percentage of births in which the mother's prenatal care began in the first trimester fails to document whether the prenatal care was adequate, nor whether it actually continued for the course of the pregnancy.

Obj. 16-9a Primary Cesarean section deliveries among low-risk females are defined as singleton, full-term, vertex presentation births among women giving birth for the first time.

Obj. 16-9b Repeat Cesarean section deliveries among low-risk females are defined as full-term, singleton, vertex presentation births among women who have had a previous Cesarean section delivery.

Obj. 16-10a,b Low birth weight (LBW) and very low birth weight (VLBW) have been associated with negative birth outcomes and as an indicator of access problems and/or need for prenatal care services. LBW is defined as the percentage of live births weighing less than 2,500 grams; VLBW is defined as the percentage of live births weighing less than 1,500 grams.

Obj. 16-11a-c Preterm births are defined as: a) less than 37 completed weeks of gestation; and b) between 32 and 36 completed weeks of gestation. Very preterm births are defined as c) less than 32 completed weeks of gestation.

AGE-ADJUSTED RATES

The numerator data used to compute rates and percentages were extracted by place of residence of the decedent for objectives using mortality data and place of residence of the mother for those using natality data.

An unstandardized rate (usually referred to as a "crude rate") is obtained by dividing the total number of vital events (e.g. deaths) by the total population at risk, then multiplying by some standard basis (e.g. per 100,000). Crude death rates, which include the effects of age, are the rates that should be applied when measuring the actual risk of dying in a specific population. Subpopulations with varying age compositions can have highly disparate death rates, since the risk of dying is primarily a function of age. Therefore, subpopulations with a large component of elderly tend to have higher death rates simply because the risk of dying is determined mostly by age. Any effect of different age compositions among population subgroups can be removed by the process known as "age-adjustment."

Age-adjusted death rates are hypothetical rates obtained by calculating age-specific rates for each subpopulation and multiplying these rates by proportions of the same age categories in a standard population, then summing the apportioned specific rates to a total. The "standard population" used in the age-adjusted death rates in this report is the 2000 United States Standard Population. The age-adjusted rates put all subgroups on an equal basis with respect to the effect of age and permit direct comparisons among subgroups and with HP 2010 mortality objectives. It is important to understand that age-adjusted death rates should be viewed as hypothetical constructs or index numbers rather than as actual measures of the risk of mortality. For further information on age-adjusted rates, see the NCHS report by Klein and Schoenborn.[11]

Natality data, prenatal care and low birth weight, were not age-adjusted and represent percentages based on the number of live births. Comparisons of infant mortality among subpopulations reflect the actual risk of dying within one year of birth and are unaffected by confounding of different age compositions because the decedents are all of the same age (under one year). Age-adjusting is not applicable to these data.

RELIABILITY OF RATES

All vital statistics rates are subject to random variation. This variation is inversely related to the number of events (e.g. death) used to calculate the rate. The smaller the frequency of occurrence of an event, then the greater the likelihood of random fluctuations within a specified time period or a certain subpopulation. The more rare an event, the less stable its occurrence is from observation to observation. As a consequence, subpopulations with only a few deaths or a few events can have highly unstable rates from year to year.

The "standard error" of a death rate and "coefficient of variation" (or Relative Standard Error, RSE) provide a rational basis for determining which rates may be considered “unreliable.” In this report, subpopulations with a relative standard error of greater than or equal to 23 percent of the rate or percent are marked with the notation “DSU” (Data Statistically Unreliable). This criterion conforms with the standard used by the NCHS in determining the reliability cut-off for rates and percents.[13] In addition, for rates based on zero events the standard error cannot be calculated and is indeterminate. Whenever the standard error is indeterminate, confidence limits are not calculated and zeroes are used to denote the confidence intervals.

TESTS OF STATISTICAL SIGNIFICANCE

A confidence interval (C.I.) is a range of values that is normally used to describe the uncertainty around a point estimate of a rate. Confidence intervals are a measure of the variability in the data and describe how much different the point estimate could have been if the underlying conditions stayed the same but chance had led to a different set of data. Confidence intervals are calculated with a stated probability (i.e., 95%), such that there is a 95% chance that the C.I. covers the true value. The true population value is a constant, even though its value is unknown, but a C.I. is a random quantity whose value depends on the random sample or data from which it is calculated. Therefore we describe a 95% C.I. as having a 95% probability of covering the true value, rather than saying that there is a 95% probability that the true value falls within the confidence interval.

When the number of events is small in relation to the population at risk (i.e., the event is rare), calculation of 95% CIs based on the Poisson probability distribution is recommended. (95% CIs correspond to a p-value of 0.05, so if you are making many comparisons approximately 5% of the comparisons may be statistically significant due to chance alone.) In general, if CIs for two separate rates overlap, there is no statistically significant difference between the two rates. In a one sample case, as for example if one is comparing the age-adjusted rate for a particular racial or ethnic group to a standard value, confidence intervals are equivalent to statistical tests. That is, if a 95% confidence interval around a particular age-adjusted rate excludes the comparison value, then a statistical test for the difference between the two values would be significant at the 0.05 level.[12]

For more detailed readings on appropriate statistical methods for comparing independent rates or percentages, please see the citations in the References section.

SUMMARY TABLE

Data for each objective are summarized in the Executive Summary, with unreliable rates or percentages noted. Included in the summary table for comparison purposes are the HP 2010 target rate, the Statewide rate (or percentage), and an indicator for whether the objective is being achieved for each gender and race-ethnic group.

FORMULAS USED IN THIS REPORT

$$IDR = \left(\frac{{}_nD}{{}_nB}\right) \times 1000$$

$$AADR = \sum W_a \left(\frac{{}_nD_a}{Npop_a}\right) \times B$$

$$ASDR = \left(\frac{{}_nD_a}{Npop_a}\right) \times B$$

$$SE_x = \left(\frac{CDR}{\sqrt{{}_nD}}\right)$$

$$SE_y = \sqrt{\sum \frac{\left(W_a \times ASDR\right)^2}{{}_nD_a}}$$

$$RSE_x = \left(\frac{SE_x}{CDR}\right) \times 100$$

$$RSE_y = \left(\frac{SE_y}{AADR}\right) \times 100$$

Lower 95% CI = AADR – (1.96 x SE_y) Upper 95% CI = AADR + (1.96 x SE_y)

Where:

IDR	= Infant Death Rate
AADR	= Age-Adjusted Death Rate
ASDR	= Age-Specific Death Rate
n^B	= Number of Live Births
${}_nD$	= Number of Deaths
Npop	= Population Size
${}_nD_a$	= Number of Deaths in an Age Group
$Npop_a$	= Population Size in Same Age Group
B	= Base (100,000)
W_a	= Age-Specific Weight (Standard Population Proportion)
SE_x	= Standard Error of a Crude Rate
RSE_x	= Relative Standard Error of a Crude Rate
SE_y	= Standard Error of an Age-Adjusted Death Rate
RSE_y	= Relative Standard Error of an Age-Adjusted Death Rate

TECHNICAL REFERENCES

1. Death Statistical Master Tape File Documentation, 2000-2002. Sacramento, CA: CDHS Center for Health Statistics, Office of Health Information and Research.
2. Documentation for the Underlying and Multiple Cause-of-Death Public Use Data Files, 2000-2001. Sacramento, CA: CDHS Center for Health Statistics, Office of Health Information and Research.
3. Birth Statistical Master Tape File Documentation, 2000-2003. Sacramento, CA: CDHS Center for Health Statistics, Office of Health Information and Research.
4. California Birth Cohort Perinatal File Documentation, 2000-2001. Sacramento, CA: CDHS Center for Health Statistics, Office of Health Information and Research.
5. Patient Discharge Data File Documentation, Public Version, 2000-2002. Sacramento, CA: Office of Statewide Health Planning and Development, Healthcare Information Resource Center.
6. California Health Interview Survey Public Use Data Dictionary, 2001. Los Angeles, CA: UCLA Data Access Center.
7. California Department of Finance. Population projections by race/ethnicity for California and its counties, 2000-2050. Sacramento, CA: Demographic Research Unit, May 2004.
8. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision. Geneva, Switzerland, WHO, 1999.
9. U.S. Department of Health and Human Services. Healthy People 2010, 2nd Edition. Washington, D.C.: U.S. Government Printing Office, November 2000.
10. U.S. Department of Health and Human Services. Tracking Healthy People 2010. Washington, D.C.: U.S. Government Printing Office, November 2000.
11. Klein, R.J. and Schoenborn, C.A. Age adjustment using the 2000 projected U.S. population. Statistical Notes, No. 20. Hyattsville, MD: National Center for Health Statistics, January 2001.
12. Armitage, P. and Berry, G. Statistical Methods in Medical Research. Boston, MA: Blackwell Scientific Publications, 1987.